ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Multivariate Analysis 98 (2007) 1684 – 1692

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Pseudo-inverse multivariate/matrix-variate distributions

Zhihua Zhang

Department of Electrical and Computer Engineering, University of University, Santa Barbara, CA, USA

Received 20 November 2005
Available online 23 May 2006

Abstract

The Moore–Penrose inverse of a singular or nonsquare matrix is not only existent but also unique. In this paper, we derive the Jacobian of the transformation from such a matrix to the transpose of its Moore–Penrose inverse. Using this Jacobian, we investigate the distribution of the Moore–Penrose inverse of a random matrix and propose the notion of pseudo-inverse multivariate/matrix-variate distributions. For arbitrary multivariate or matrix-variate distributions, we can develop the corresponding pseudo-inverse distributions. In particular, we present pseudo-inverse multivariate normal distributions, pseudo-inverse Dirichlet distributions, pseudo-inverse matrix-variate normal distributions and pseudo-inverse Wishart distributions.

AMS 1991 subject classification: 62H10; 62E15; 15A09; 15A52

Keywords: Matrix-variate distribution; The Moore–Penrose generalized inverse; Pseudo-inverse multivariate distribution; Pseudo-inverse matrix-variate distribution

1. Introduction

We are concerned with the problem of random matrices and their distributions. Let $\mathbf{X}$ be an $n \times n$ nonsingular random matrix and denote its inverse by $\mathbf{X}^{-1}$. It is direct to obtain the distribution of $\mathbf{X}^{-1}$ from the distribution of $\mathbf{X}$ by using the Jacobian of the transformation from $\mathbf{X}$ to $\mathbf{X}^{-1}$. For example, we can derive an inverted Wishart distribution from the corresponding Wishart distribution [5]. However, we usually meet the case that either $\mathbf{X}$ is singular or it is not square, such as an $n \times n$ singular Wishart matrix $\mathbf{X}$ [10,1,4,8] or an $n \times m$ matrix-variate normal matrix $\mathbf{X}$ [5]. In this case, it is impossible to derive the inverted distribution of $\mathbf{X}$ because its inverse does not exist.

E-mail address: zhzhang@ece.ucsb.edu.

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.04.002

It is well known that the Moore–Penrose generalized inverse of $\mathbf{X}$ exists and is unique [7]. In this paper, motivated by the existence and uniqueness of the Moore–Penrose inverse, we discuss the distribution of the Moore–Penrose inverse of a random matrix. Consequently, we develop a family of multivariate/matrix-variate distributions, which are called *pseudo-inverse multivariate/matrix-variate distributions*.

In Section 2, we consider a nonzero $n\times 1$ matrix, i.e., a nonzero n-dimensional random vector $\mathbf{x}$. For this special case, we present an explicit transformation from $\mathbf{x}$ to the transpose of its Moore–Penrose inverse so that the Jacobian of this transformation is easily calculated. Accordingly, given an arbitrary nonzero random multivariate vector, we can obtain its pseudo-inverse distribution. We present two specific pseudo-inverse multivariate distributions: pseudo-inverse multivariate normal distribution and pseudo-inverse Dirichlet distribution. Given an $n\times m$ matrix $\mathbf{X}$, we in Section 3 derive a generic Jacobian of the transformation from $\mathbf{X}$ to the transpose of its Moore–Penrose inverse. When $\mathbf{X}$ is square yet symmetric, the Jacobian is also given. In Section 4, we propose a notion of pseudo-inverse matrix-variate distribution. In particular, we give two distributions: pseudo-inverse matrix-variate normal distribution and pseudo-inverse Wishart distribution. The latter is a generalization of the inverted Wishart distribution. We prove that there exists the dual relationship between a pseudo-inverse matrix-variate normal distribution and an inverted Wishart distribution. Finally, we give some conclusions in Section 5.

2. Pseudo-inverse multivariate distributions

First of all, we present some notations. Let $\mathbf{I}_m$ denote the $m\times m$ identity matrix and let $\mathbf{0}$ denote the zero vector (or matrix) whose dimensionality is dependent upon the context. We write $\mathbf{A}\succ 0$ if $\mathbf{A}$ is positive definite and $\mathbf{A}\succeq 0$ if $\mathbf{A}$ is positive semidefinite. In addition, $\mathbf{A}\otimes\mathbf{B}$ represents the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$.

Let $\mathbf{X}$ be an $n\times m$ matrix. An $m\times n$ matrix $\mathbf{X}^+$ is said to be the Moore–Penrose inverse of $\mathbf{X}$ if conditions

$$\text{(i)}\quad \mathbf{X}\mathbf{X}^+\mathbf{X}=\mathbf{X},\qquad \text{(iii)}\quad (\mathbf{X}\mathbf{X}^+)'=\mathbf{X}\mathbf{X}^+,$$

$$\text{(ii)}\quad \mathbf{X}^+\mathbf{X}\mathbf{X}^+=\mathbf{X}^+,\qquad \text{(iv)}\quad (\mathbf{X}^+\mathbf{X})'=\mathbf{X}^+\mathbf{X}$$

are satisfied. The following properties will be useful for our subsequent studies. Interested readers are referred to [7, p. 34–36] for more details.

Proposition 1. *Let* $\mathbf{X}^+$ $(m\times n)$ *be the Moore–Penrose inverse of* $\mathbf{X}$ $(n\times m)$. *Then*

(1) $\mathbf{X}^+$ *exists and is unique*.
(2) $(\mathbf{X}^+)^+=\mathbf{X}$.
(3) $\mathbf{X}^+=\mathbf{0}$ *if and only if* $\mathbf{X}=\mathbf{0}$.
(4) $\mathbf{X}^+$ *and* $\mathbf{X}$ *have the same rank*.
(5) $(\mathbf{X}\mathbf{X}')^+=(\mathbf{X}^+)'\mathbf{X}^+$ *and* $(\mathbf{X}'\mathbf{X})^+=\mathbf{X}^+(\mathbf{X}^+)'$.
(6) $\mathbf{X}=\mathbf{U}'\begin{pmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{V}$ *is the singular value decomposition* (*SVD*) *of* $\mathbf{X}$, *then*

$$\mathbf{X}^+=\mathbf{V}'\begin{pmatrix}\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{U}.$$

Throughout this paper, we assume that $\mathbf{X} \neq \mathbf{0}$ due to Proposition 1(3). In this section, we consider a simple case. We are given an n-dimensional nonzero random vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$. Let $\mathbf{y}' = (y_1, y_2, \ldots, y_n)$ be the Moore–Penrose inverse of $\mathbf{x}$. It is easy to verify that $\mathbf{y} = \frac{1}{\|\mathbf{x}\|^2}\mathbf{x}$ where $\|\mathbf{x}\|^2 = \sum_{j=1}^n x_j^2$. Further, we have

$$\mathbf{x} = \frac{1}{\|\mathbf{y}\|^2}\mathbf{y} \tag{1}$$

because of $\|\mathbf{x}\|^2 = \frac{1}{\|\mathbf{y}\|^2}$.

The Jacobian matrix of $\mathbf{y}$ with respect to (w.r.t.) $\mathbf{x}$ is $\frac{\partial \mathbf{y}}{\partial \mathbf{x}'} = [\partial y_i / \partial x_j]$ $(n \times n)$ where

$$\frac{\partial y_i}{\partial x_j} = \begin{cases} \frac{1}{\|\mathbf{x}\|^2} - \frac{2}{\|\mathbf{x}\|^4} x_i x_i = \|\mathbf{y}\|^2 - 2y_i y_j, & i = j, \\ -\frac{2}{\|\mathbf{x}\|^4} x_i x_j = -2y_i y_j, & i \neq j. \end{cases}$$

Hence, $\frac{\partial \mathbf{y}}{\partial \mathbf{x}'} = \|\mathbf{y}\|^2 \mathbf{I}_n - 2\mathbf{y}\mathbf{y}'$. Since the eigenvalues of $\mathbf{y}\mathbf{y}'$ are 0 with multiplicity $(n-1)$ and $\|\mathbf{y}\|^2$, the eigenvalues of $\frac{\partial \mathbf{y}}{\partial \mathbf{x}'}$ are $\|\mathbf{y}\|^2$ with multiplicity $(n-1)$ and $-\|\mathbf{y}\|^2$. The Jacobian is thus given by

$$J(\mathbf{y} \to \mathbf{x}) = \mathrm{mod} \left| \frac{\partial \mathbf{y}}{\partial \mathbf{x}'} \right| = \|\mathbf{y}\|^{2n} = \|\mathbf{x}\|^{-2n}. \tag{2}$$

Now given a distribution of $\mathbf{x}$, we can immediately obtain the corresponding distribution of $\mathbf{y}$ using (1) and (2). We refer to the distribution of $\mathbf{y}$ as *pseudo-inverse multivariate distribution* w.r.t. $\mathbf{x}$. It is obvious that for arbitrary existing multivariate distributions, we can derive their corresponding pseudo-inverse multivariate distributions. Here we only take two examples to illustrate the constructions of pseudo-inverse multivariate distributions. First, if $\mathbf{x} \sim \mathcal{N}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $\mathbf{y}$ follows a pseudo-inverse multivariate normal distribution and the p.d.f. is given by

$$(2\pi)^{-\frac{n}{2}} \|\mathbf{y}\|^{-2n} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\|\mathbf{y}\|^4} \left(\mathbf{y} - \|\mathbf{y}\|^2 \boldsymbol{\mu}\right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{y} - \|\mathbf{y}\|^2 \boldsymbol{\mu}\right) \right\}.$$

Second, let $\mathbf{x}$ follow the Dirichlet distribution with density

$$\frac{\Gamma(\sum_{i=0}^n \alpha_i)}{\prod_{i=0}^n \Gamma(\alpha_i)} \left(1 - \sum_{i=1}^n x_i\right)^{\alpha_0 - 1} \prod_{i=1}^n x_i^{\alpha_i - 1}, \quad 0 < x_i < 1, \ \sum_{i=1}^n x_i \leqslant 1,$$

where $\alpha_i > 0$, $i = 0, 1, \ldots, n$, are parameters and $\Gamma(\cdot)$ is the Gamma function. Then $\mathbf{y}$ follows a pseudo-inverse Dirichlet distribution and the p.d.f. is given by

$$\frac{\Gamma(\sum_{i=0}^n \alpha_i)}{\prod_{i=0}^n \Gamma(\alpha_i)} \|\mathbf{y}\|^{2(1-\sum_{i=0}^n \alpha_i)} \left(\|\mathbf{y}\|^2 - \sum_{i=1}^n y_i\right)^{\alpha_0 - 1} \prod_{i=1}^n y_i^{\alpha_i - 1}, \quad 0 < y_i < \|\mathbf{y}\|^2,$$
$$\sum_{i=1}^n y_i \leqslant \|\mathbf{y}\|^2.$$

Notice that our pseudo-inverse multivariate normal distribution and pseudo-inverse Dirichlet distribution differ from the inverse Gaussian (Wald) distribution and the inverted Dirichlet [6].

3. The Jacobian of the transformation from $(\mathbf{X}^+)'$ to $\mathbf{X}$

We are now given an $n\times m$ real nonzero random matrix $\mathbf{X}$. As can be seen from the previous section, $\mathbf{X}^+$ is $m\times n$. Since $\mathbf{X}$ and $(\mathbf{X}^+)'$ have the same size, this motivates us to calculate the Jacobian $J((\mathbf{X}^+)' \to \mathbf{X})$, rather than $J(\mathbf{X}^+ \to \mathbf{X})$. We denote $\mathbf{Y} = (\mathbf{X}^+)'$ $(n\times m)$ and refer to the distribution of $\mathbf{Y}$ as *pseudo-inverse matrix-variate distribution* w.r.t. $\mathbf{X}$.

In order to derive the distribution of $\mathbf{Y}$ from that of $\mathbf{X}$, it is necessary to obtain the Jacobian $J(\mathbf{Y} \to \mathbf{X})$. Assume that $\mathbf{X}$ is of rank r with $r \leqslant \min(n, m)$. Proposition 1(4) shows that the rank of $\mathbf{Y}$ is also r. Moreover, it is clear that the rank of $\mathbf{XX}'$ (**or** $\mathbf{YY}'$) is r and $\mathbf{XX}'$ (**or** $\mathbf{YY}'$) is positive semidefinite. In addition, $\mathbf{XX}'$ (**or** $\mathbf{YY}'$) as well $\mathbf{X}'\mathbf{X}$ (**or** $\mathbf{Y}'\mathbf{Y}$) has r positive eigenvalues because both have the same nonzero eigenvalues.

Lemma 2. *Let $\mathbf{Y}'$ be the Moore–Penrose inverse of an $n\times m$ real matrix $\mathbf{X}$. Then the Jacobian of the transformation from $\mathbf{Y}$ to $\mathbf{X}$ is given by*

$$J(\mathbf{Y} \to \mathbf{X}) = \prod_{i=1}^{r} \lambda_i^{(n+m-r)}, \tag{3}$$

where $r \leqslant \min(n, m)$ is the rank of $\mathbf{X}$ (or $\mathbf{Y}$), and λ_i and λ_i^{-1} are, respectively, the positive eigenvalues of $\mathbf{YY}'$ and $\mathbf{XX}'$, for $i = 1, \ldots, r$.

Remarks. Without loss of generality, we assume $n \geqslant m$. If $\mathbf{X}$ is of full rank, i.e., $r{=}m$, then $J(\mathbf{Y} \to \mathbf{X}) = |\mathbf{X}'\mathbf{X}|^{-n} = |\mathbf{Y}'\mathbf{Y}|^{n}$. Further, when $m = 1$, Eq. (3) reduces to Eq. (2). If $\mathbf{X}$ is square and of full rank, i.e., $n{=}m{=}r$, then $J(\mathbf{Y} \to \mathbf{X}) = |\mathbf{X}|^{-2n}$, a well-known result (see, [5, p. 14]).

Proof. Using Proposition 1(6), we have

$$\mathbf{X} = \mathbf{U}' \begin{array}{c} \\ r \\ n-r \end{array} \begin{array}{c} \begin{array}{cc} r & \quad m-r \end{array} \\ \begin{pmatrix} D & \quad 0 \\ 0 & \quad 0 \end{pmatrix} \end{array} \mathbf{V} \quad \text{and} \quad \mathbf{Y} = \mathbf{U}' \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V}, \tag{4}$$

where $\mathbf{D}$ $(r\times r)$ is a diagonal matrix with positive elements, and $\mathbf{U}$ $(n\times n)$ and $\mathbf{V}$ $(m\times m)$ satisfy

$$\mathbf{UU}' = \mathbf{I}_n \quad \text{and} \quad \mathbf{VV}' = \mathbf{I}_m. \tag{5}$$

Accordingly, we have

$$d\mathbf{X} = (d\mathbf{U}') \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V} + \mathbf{U}' \begin{pmatrix} d\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V} + \mathbf{U}' \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (d\mathbf{V})$$

and

$$d\mathbf{Y} = (d\mathbf{U}') \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V} - \mathbf{U}' \begin{pmatrix} \mathbf{D}^{-1}(d\mathbf{D})\mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V} + \mathbf{U}' \begin{pmatrix} \mathbf{D}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (d\mathbf{V}),$$

which lead to

$$\mathbf{U}(d\mathbf{X})\mathbf{V}' = \mathbf{U}(d\mathbf{U}') \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} d\mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (d\mathbf{V})\mathbf{V}'$$

and

$$\mathbf{U}(d\mathbf{Y})\mathbf{V}' = \mathbf{U}(d\mathbf{U}')\begin{pmatrix}\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} - \begin{pmatrix}\mathbf{D}^{-1}(d\mathbf{D})\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} + \begin{pmatrix}\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}(d\mathbf{V})\mathbf{V}'.$$

Let $\mathbf{G} = \mathbf{U}(d\mathbf{U}')$ $(n\times n)$ and $\mathbf{H} = (d\mathbf{V})\mathbf{V}'$ $(m\times m)$. Partitioning them as

$$\mathbf{G} = \begin{array}{c c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ n-r \end{array} & \begin{pmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{pmatrix} \end{array}, \quad \mathbf{H} = \begin{array}{c c} & \begin{array}{cc} r & m-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{pmatrix}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{pmatrix} \end{array},$$

we then obtain

$$\begin{aligned}
\mathbf{U}(d\mathbf{X})\mathbf{V}' &= \begin{pmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} + \begin{pmatrix}d\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} + \begin{pmatrix}\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{pmatrix}\\
&= \begin{pmatrix}\mathbf{G}_{11}\mathbf{D} + d\mathbf{D} + \mathbf{D}\mathbf{H}_{11} & \mathbf{D}\mathbf{H}_{12}\\ \mathbf{G}_{21}\mathbf{D} & \mathbf{0}\end{pmatrix}\\
&\triangleq \begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix} = \mathbf{A} \quad \text{(denoted)}
\end{aligned} \tag{6}$$

and

$$\begin{aligned}
&\mathbf{U}(d\mathbf{Y})\mathbf{V}'\\
&= \begin{pmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} - \begin{pmatrix}\mathbf{D}^{-1}(d\mathbf{D})\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} + \begin{pmatrix}\mathbf{D}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{H}_{11} & \mathbf{H}_{12}\\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{pmatrix}\\
&= \begin{pmatrix}\mathbf{G}_{11}\mathbf{D}^{-1} - \mathbf{D}^{-1}(d\mathbf{D})\mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{H}_{11} & \mathbf{D}^{-1}\mathbf{H}_{12}\\ \mathbf{G}_{21}\mathbf{D}^{-1} & \mathbf{0}\end{pmatrix}\\
&\triangleq \begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix} = \mathbf{B} \quad \text{(denoted).}
\end{aligned} \tag{7}$$

On the other hand, it follows from (5) that

$$\mathbf{G} = -(d\mathbf{U})\mathbf{U}' = -\mathbf{G}' \quad \text{and} \quad \mathbf{H} = -\mathbf{V}(d\mathbf{V}') = -\mathbf{H}'.$$

We thus have $\mathbf{G}_{11} = -\mathbf{G}'_{11}$ and $\mathbf{H}_{11} = -\mathbf{H}'_{11}$. Again from (6) and (7), it is easy to obtain

$$\begin{aligned}
&\mathbf{B}_{11} = -\mathbf{D}^{-1}\mathbf{A}'_{11}\mathbf{D}^{-1},\\
&\mathbf{B}_{12} = \mathbf{D}^{-2}\mathbf{A}_{12},\\
&\mathbf{B}_{21} = \mathbf{A}_{21}\mathbf{D}^{-2},\\
&\mathbf{B}_{22} = \mathbf{A}_{22} = \mathbf{0}.
\end{aligned}$$

Hence,

$$\begin{aligned}
J(\mathbf{Y} \to \mathbf{X}) &= J(d\mathbf{Y} \to d\mathbf{X})\\
&= J(d\mathbf{Y} \to \mathbf{B})J(\mathbf{B}_{11}, \mathbf{B}_{12}, \mathbf{B}_{21} \to \mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21})J(\mathbf{A} \to d\mathbf{X})\\
&= J(\mathbf{B}_{11} \to \mathbf{A}_{11})J(\mathbf{B}_{12} \to \mathbf{A}_{12})J(\mathbf{B}_{21} \to \mathbf{A}_{21})\\
&= J(\mathbf{B}_{11} \to \mathbf{A}'_{11})J(\mathbf{A}'_{11} \to \mathbf{A}_{11})J(\mathbf{B}_{12} \to \mathbf{A}_{12})J(\mathbf{B}_{21} \to \mathbf{A}_{21})\\
&= |\mathbf{D}|^{-2r}|\mathbf{D}|^{-2(n-r)}|\mathbf{D}|^{-2(m-r)}\\
&= |\mathbf{D}|^{-2(n+m-r)}.
\end{aligned}$$

Since we can write $\mathbf{XX}' = \mathbf{U}_1'\mathbf{D}^2\mathbf{U}_1$ and $\mathbf{YY}' = \mathbf{U}_1'\mathbf{D}^{-2}\mathbf{U}_1$, where $\mathbf{U}_1$ $(r\times n)$ consists of the first r rows of $\mathbf{U}$ subject to $\mathbf{U}_1\mathbf{U}_1' = \mathbf{I}_r$, the diagonal elements (say, λ_i^{-1}'s) of $\mathbf{D}^2$ are the r nonzero eigenvalues of $\mathbf{XX}'$ while λ_i, $i = 1, \ldots, r$, are the nonzero eigenvalues of $\mathbf{YY}'$. Thus we obtain (3). □

Now we consider case that $\mathbf{X}$ is an $n\times n$ real symmetric matrix with rank r $(\leqslant n)$. In this case, we rewrite (6) and (7) as

$$\mathbf{U}(d\mathbf{X})\mathbf{U}' = \begin{pmatrix} \mathbf{G}_{11}\mathbf{D} + d\mathbf{D} - \mathbf{DG}_{11} & -\mathbf{DG}_{12} \\ \mathbf{G}_{21}\mathbf{D} & \mathbf{0} \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} = \mathbf{A} \tag{8}$$

and

$$\mathbf{U}(d\mathbf{Y})\mathbf{U}' = \begin{pmatrix} \mathbf{G}_{11}\mathbf{D}^{-1} - \mathbf{D}^{-1}(d\mathbf{D})\mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{G}_{11} & -\mathbf{D}^{-1}\mathbf{G}_{12} \\ \mathbf{G}_{21}\mathbf{D}^{-1} & \mathbf{0} \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \mathbf{B} \tag{9}$$

due to $\mathbf{U} = \mathbf{V}$ and $\mathbf{H} = -\mathbf{G}$. Thus,

$$\mathbf{B}_{11} = -\mathbf{D}^{-1}\mathbf{A}_{11}\mathbf{D}^{-1},$$
$$\mathbf{B}_{12} = \mathbf{B}_{21}' = \mathbf{D}^{-2}\mathbf{A}_{12},$$
$$\mathbf{B}_{22} = \mathbf{A}_{22} = \mathbf{0}.$$

and hence,

$$\begin{aligned} &J(\mathbf{Y} \to \mathbf{X}) \\ &\quad = J(d\mathbf{Y} \to d\mathbf{X}) \\ &\quad = J(d\mathbf{Y} \to \mathbf{B}_{11}, \mathbf{B}_{12})J(\mathbf{B}_{11}, \mathbf{B}_{12} \to \mathbf{A}_{11}, \mathbf{A}_{12})J(\mathbf{A}_{11}, \mathbf{A}_{12} \to d\mathbf{X}) \\ &\quad = J(\mathbf{B}_{11} \to \mathbf{A}_{11})J(\mathbf{B}_{12} \to \mathbf{A}_{12}) = |\mathbf{D}|^{-r-1}|\mathbf{D}|^{-2(n-r)} = |\mathbf{D}|^{-2n+r-1}. \end{aligned}$$

In addition, it is clearly seen that the diagonal elements of $\mathbf{D}$ (resp., $\mathbf{D}^{-1}$) are the nonzero eigenvalues of $\mathbf{X}$ (resp., $\mathbf{Y}$). Therefore, we have the following Jacobian.

Lemma 3. *Let* $\mathbf{Y}$ *be the Moore–Penrose inverse of the* $n\times n$ *real symmetric matrix* $\mathbf{X}$ *and the rank of* $\mathbf{X}$ *be* r, *then*

$$J(\mathbf{Y} \to \mathbf{X}) = \prod_{i=1}^{r} |\eta_i|^{2n-r+1}, \tag{10}$$

where η_i *(resp.,* η_i^{-1}*),* $i = 1, \ldots, r$, *are the* r *nonzero eigenvalues of* $\mathbf{Y}$ *(resp.,* $\mathbf{X}$*).*

Remark. If $\mathbf{X}$ is symmetric and nonsingular, i.e., $r{=}n$, then $J(\mathbf{Y} \to \mathbf{X}) = |\mathbf{X}|^{-(n+1)} = |\mathbf{Y}|^{n+1}$ [5, p.14]. We note that the similar results as our Lemmas 2 and 3 have been independently proposed in [2,3]. The main difference is in that the derivations of [2,3] are based on exterior products, whereas ours based on the singular value decomposition. In addition, since Theorem 1 of [3] requires that the matrices involved are positive semidefinite, this theorem is a special case of Lemma 3. More importantly, in the results of [2,3], the nonzero eigenvalues of the matrix involved are assumed to be mutually different. However, this assumption is no longer necessary in our lemmas.

4. Pseudo-inverse matrix-variate distributions

Equipped with an arbitrary matrix-variate distribution, it is ready to derive the corresponding pseudo-inverse distribution from Lemmas 2 or 3. In this section, given two spacial matrix-variate distributions: the matrix-variate normal distribution and the singular Wishart distribution, we study the pseudo-inverse matrix-variate distributions w.r.t. them.

Consider an $n \times m$ random matrix $\mathbf{X}$, which follows the matrix-variate normal distribution with mean matrix $\mathbf{M}$ and covariance matrix $\mathbf{\Sigma} \otimes \mathbf{\Psi}$ where $\mathbf{\Sigma}(n \times n) \succ 0$ and $\mathbf{\Psi}(m \times m) \succ 0$. The p.d.f. of $\mathbf{X}$ is given by

$$p(\mathbf{X}) = (2\pi)^{-nm/2} |\mathbf{\Sigma}|^{-m/2} |\mathbf{\Psi}|^{-n/2} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\mathbf{\Sigma}^{-1} (\mathbf{X} - \mathbf{M}) \mathbf{\Psi}^{-1} (\mathbf{X} - \mathbf{M})' \right) \right]. \tag{11}$$

We use the notation $\mathbf{X} \sim \mathcal{N}_{n,m}(\mathbf{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$. We present the following definition.

Definition 4. If $\mathbf{X} \sim \mathcal{N}_{n,m}(\mathbf{M}, \mathbf{\Sigma} \otimes \mathbf{\Psi})$, then $\mathbf{Y} = (\mathbf{X}^{+})'$ is said to have a pseudo-inverse matrix-variate normal distribution, written as $\mathbf{Y} \sim \mathcal{IN}_{n,m}(\mathbf{M}, \mathbf{\Sigma}^{-1} \otimes \mathbf{\Psi}^{-1})$.

According to Definition 4 and Lemma 2, we have the p.d.f. of $\mathbf{Y}$:

$$p(\mathbf{Y}) = \frac{\prod_{i=1}^{r} \lambda_i^{-(n+m-r)}}{(2\pi)^{nm/2} |\mathbf{\Sigma}|^{m/2} |\mathbf{\Psi}|^{n/2}} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\mathbf{\Sigma}^{-1} ((\mathbf{Y}^{+})' - \mathbf{M}) \mathbf{\Psi}^{-1} (\mathbf{Y}^{+} - \mathbf{M}') \right) \right], \tag{12}$$

where λ_i $(i = 1, \ldots, r)$ are the r nonzero eigenvalues of $\mathbf{Y}\mathbf{Y}'$.

We are usually interested in the case that $\mathbf{\Psi} = \mathbf{I}_m$. Without loss of generality, we also let $\mathbf{M} = \mathbf{0}$. Thus, the columns of $\mathbf{X}$ are independent from $\mathcal{N}_n(\mathbf{0}, \mathbf{\Sigma})$ and $\mathbf{X}$ is of full rank with probability one [9, p. 73], implying that the rank of $\mathbf{Y}$ is equal to $\min(n, m)$ with probability one. However, the columns of $\mathbf{Y}$ are not independent of each other.

If $m \geqslant n$, the p.d.f. of $\mathbf{Y}$ becomes

$$p(\mathbf{Y}) = \frac{|\mathbf{Y}\mathbf{Y}'|^{-m}}{(2\pi)^{nm/2} |\mathbf{\Sigma}|^{m/2}} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\mathbf{\Sigma}^{-1} (\mathbf{Y}^{+})' \mathbf{Y}^{+} \right) \right]$$

due to $r = n$ and $\prod_{i=1}^{r} \lambda_i = |\mathbf{Y}\mathbf{Y}'|$. Further, $\mathbf{K} = \mathbf{X}\mathbf{X}'$ is nonsingular with probability one and it follows Wishart distribution $\mathcal{W}_n(m, \ \mathbf{\Sigma})$ [5, p. 88]. Since the inverse $\mathbf{K}^{-1}$ of $\mathbf{K}$ exists, $\mathbf{K}^{-1}$ is distributed according to inverted Wishart distribution $\mathcal{IW}_n(m, \ \mathbf{\Sigma}^{-1})$.[2]

We now consider the case of $m < n$. The ranks of both $\mathbf{Y}$ and $\mathbf{Y}'\mathbf{Y}$ are m. Thus, we can write the p.d.f. of $\mathbf{Y}$ as

$$p(\mathbf{Y}) = \frac{|\mathbf{Y}'\mathbf{Y}|^{-n}}{(2\pi)^{nm/2} |\mathbf{\Sigma}|^{m/2}} \exp\left[-\frac{1}{2} \mathrm{tr}\left(\mathbf{\Sigma}^{-1} (\mathbf{Y}^{+})' \mathbf{Y}^{+} \right) \right].$$

Moreover, $\mathbf{K}$ is singular and its rank is equal to m with probability one. In this case, $\mathbf{K}$ is said to have a singular Wishart distribution. Recently, Srivastava [8] presented the p.d.f. of the singular Wishart distribution. Let $\mathbf{W} = \mathbf{K}^{+}$ be the Moore–Penrose inverse of $\mathbf{K}$. We then say $\mathbf{W}$ to have a pseudo-inverse Wishart distribution. It follows from Proposition 1(5) that $\mathbf{W} = \mathbf{Y}\mathbf{Y}'$, and the

[2] It is denoted by $\mathcal{IW}_n(m+n+1, \ \mathbf{\Sigma}^{-1})$ in [5, p. 111]. In contrast to $\mathcal{W}_n(m, \ \mathbf{\Sigma})$, we here use $\mathcal{IW}_n(m, \ \mathbf{\Sigma}^{-1})$.

rank of $\mathbf{W}$ is m $(< n)$. Hence, there always exists such a permutation matrix (say, $\mathbf{P}$) that the first $m \times m$ principal submatrix of $\mathbf{PWP}'$ is nonsingular. Without loss of generality, we then partition $\mathbf{W}$ as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix},$$

where $\mathbf{W}_{11}$ is an $m \times m$ nonsingular matrix, $\mathbf{W}_{12} = \mathbf{W}'_{21}$ and $\mathbf{W}_{22} = \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12}$. From our Lemma 3 and Theorem 3.1 in Srivastava [8], we obtain the p.d.f. of $\mathbf{W}$.

Theorem 5. *Let* $\mathbf{W}$ *be an* $n \times n$ *symmetric positive semidefinite matrix with rank* m $(< n)$. *If* $\mathbf{W}$ *is the pseudo-inverse Wishart distribution*, *then the joint density of* $\mathbf{W}_{11}$ *and* $\mathbf{W}_{12}$ *is given by*

$$\frac{\pi^{m(m-n)/2}}{2^{nm/2}\Gamma_m(\frac{m}{2})}|\mathbf{\Sigma}|^{-m/2}|\mathbf{W}_{11}|^{-(m+n+1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{W}^{+})\right), \tag{13}$$

and $\mathbf{W}_{22} = \mathbf{W}_{21}\mathbf{W}_{11}^{-1}\mathbf{W}_{12}$. *Here* $\Gamma_m(\frac{m}{2}) = \pi^{m(m-1)/4}\prod_{j=1}^{m}\Gamma(\frac{m+1-j}{2})$ *is a normalization term.*

In the rest of this paper, we shall not strictly distinguish the singular Wishart distribution and the pseudo-inverse Wishart distribution from the Wishart distribution and the inverted Wishart distribution. We shall still use the notations $\mathbf{K} \sim \mathcal{W}_n(m,\ \mathbf{\Sigma})$ and $\mathbf{W} \sim \mathcal{IW}_n(m,\ \mathbf{\Sigma}^{-1})$ for the singular Wishart distribution and the pseudo-inverse Wishart distribution. Notice that Theorem 3.2.2 in [5] and Theorem 3.1 in [8] show that if $\mathbf{X} \sim \mathcal{N}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}\otimes\mathbf{I}_m)$, then $\mathbf{K} = \mathbf{XX}' \sim \mathcal{W}_n(m,\ \mathbf{\Sigma})$. Theorem 3.3.3 in [5] shows that if $\mathbf{K} \sim \mathcal{W}_n(m,\ \mathbf{\Sigma})$ and $m \geqslant n$ is an integer, then $\mathbf{K} = \mathbf{XX}'$ and $\mathbf{X} \sim \mathcal{N}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}\otimes\mathbf{I}_m)$. In fact, according to the proof method used in Theorem 3.3.3, it is easy to extend Theorem 3.3.3 to one where m is an integer less than n. Therefore, there exists the dual relationship between a matrix-variate normal distribution and a Wishart distribution. Interestingly, we can also obtain such duality between a pseudo-inverse matrix-variate normal distribution and an inverted Wishart distribution. That is,

Theorem 6. *Let* $\mathbf{Y} \sim \mathcal{IN}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}^{-1}\otimes\mathbf{I}_m)$, *then* $\mathbf{W} = \mathbf{YY}' \sim \mathcal{IW}_n(m,\ \mathbf{\Sigma}^{-1})$. *Conversely*, *let* $\mathbf{W} \sim \mathcal{IW}_n(m,\ \mathbf{\Sigma}^{-1})$, *and* m *be an integer*, *then* $\mathbf{W} = \mathbf{YY}'$ *and* $\mathbf{Y} \sim \mathcal{IN}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}^{-1}\otimes\mathbf{I}_m)$.

Proof. The proof is simple. First, if $\mathbf{Y} \sim \mathcal{IN}_{n,m}(\mathbf{0}, \mathbf{\Sigma}^{-1}\otimes\mathbf{I}_m)$, then $\mathbf{X} = (\mathbf{Y}^{+})' \sim \mathcal{N}_{n,m}(\mathbf{0}, \mathbf{\Sigma}\otimes\mathbf{I}_m)$. Hence, $\mathbf{K} = \mathbf{XX}' \sim \mathcal{W}_n(m,\ \mathbf{\Sigma})$. Since $\mathbf{K}^{+} = (\mathbf{XX}')^{+} = (\mathbf{X}^{+})'\mathbf{X}^{+} = \mathbf{YY}' = \mathbf{W}$, it then follows that $\mathbf{W} \sim \mathcal{IW}_n(m,\ \mathbf{\Sigma}^{-1})$. Second, let $\mathbf{W} \sim \mathcal{IW}_n(m,\ \mathbf{\Sigma}^{-1})$ and $\mathbf{K} = \mathbf{W}^{+}$. Then $\mathbf{K} \sim \mathcal{W}_n(m,\ \mathbf{\Sigma})$. There exists an $n \times m$ matrix (say, $\mathbf{X}$) such that $\mathbf{XX}' = \mathbf{K}$ and $\mathbf{X} \sim \mathcal{N}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}\otimes\mathbf{I}_m)$. Now letting $\mathbf{Y} = (\mathbf{X}^{+})'$, we have $\mathbf{Y} \sim \mathcal{IN}_{n,m}(\mathbf{0},\ \mathbf{\Sigma}^{-1}\otimes\mathbf{I}_m)$. Moreover, we obtain $\mathbf{YY}' = \mathbf{W}$ due to $\mathbf{W} = \mathbf{K}^{+} = (\mathbf{X}^{+})'\mathbf{X}^{+} = \mathbf{YY}'$. $\square$

5. Conclusion

In this paper, we have proposed a notion of pseudo-inverse distribution to model a singular or nonsquare random matrix. Our departure point came from the fact that the Moore–Penrose inverse of any matrix exists and is unique. We derived the Jacobian of the transformation between a matrix and the transpose of its Moore–Penrose inverse, and then produced an approach to the constructions of pseudo-inverse distributions. In particular, we illustrated four pseudo-inverse distributions: pseudo-inverse multivariate normal distribution, pseudo-inverse Dirichlet distribution, pseudo-inverse matrix-variate normal distribution and pseudo-inverse Wishart distribution. We have presented the duality between the pseudo-inverse matrix-variate normal distribution and the inverted Wishart distribution.

References

[1] J.A. Díaz-García, R. Gutiérrez-Jáimez, Proof of the conjectures of H. Uhlig on the singular multivariate beta and the Jacobian of a certain matrix transformation, Ann. Statist. 25 (1997) 2018–2023.
[2] J.A. Díaz-García, R. Gutiérrez-Jáimez, Functions of singular random matrices with applications, Test 14 (2005) 475–487.
[3] J.A. Díaz-García, R. Gutiérrez-Jáimez, Distribution of generalized inverse of a random matrix and its applications, J. Statist. Plann. Inference 136 (2006) 183–192.
[4] J.A. Díaz-García, R. Gutiérrez-Jáimez, K.V. Mardia, Wishart and pseudo-Wishart distributions and some applications to shape theory, J. Multivariate Anal. 63 (1997) 73–87.
[5] A.K. Gupta, D.K. Nagar, Matrix Variate Distributions, Chapman & Hall, CRC, London, Boca Raton, FL, 2000.
[6] N.L. Johnson, S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, Wiley, New York, 1972.
[7] H. Lütkepohl, Handbook of Matrices, Wiley, New York, 1996.
[8] M.S. Srivastava, Singular Wishart and multivariate Beta distributions, Ann. Statist. 31 (5) (2003) 1537–1560.
[9] M.S. Srivastava, C.G. Khatri, An Introduction to Multivariate Statistics, Elsevier North Holland, New York, 1979.
[10] H. Uhlig, On singular Wishart and singular multivariate beta distributions, Ann. Statist. 22 (1994) 395–405.